REVIEW

Korean J Intern Med 2018;33:277-283
https://doi.org/10.3904/kjim.2016.195

An overview of meta-analysis for clinicians

Young Ho Lee

Division of Rheumatology, Department of Internal Medicine, Korea University Medical Center, Seoul, Korea

Received: June 27, 2016
Accepted: September 5, 2017

Correspondence to
Young Ho Lee, M.D.
Division of Rheumatology, Department of Internal Medicine, Korea University Anam Hospital, 73 Inchon-ro, Seongbuk-gu, Seoul 02841, Korea
Tel: +82-2-920-5645
Fax: +82-2-922-5974
E-mail: lyhcgh@korea.ac.kr

The number of medical studies being published is increasing exponentially, and clinicians must routinely process large amounts of new information. Moreover, the results of individual studies are often insufficient to provide confident answers, as their results are not consistently reproducible. A meta-analysis is a statistical method for combining the results of different studies on the same topic and it may resolve conflicts among studies. Meta-analysis is being used increasingly and plays an important role in medical research. This review introduces the basic concepts, steps, advantages, and caveats of meta-analysis, to help clinicians understand it in clinical practice and research. A major advantage of a meta-analysis is that it produces a precise estimate of the effect size, with considerably increased statistical power, which is important when the power of the primary study is limited because of a small sample size. A meta-analysis may yield conclusive results when individual studies are inconclusive. Furthermore, meta-analyses investigate the source of variation and different effects among subgroups. In summary, a meta-analysis is an objective, quantitative method that provides less biased estimates on a specific topic. Understanding how to conduct a meta-analysis aids clinicians in the process of making clinical decisions.

Keywords: Meta-analysis; Statistical analysis; Advantage; Limitation

INTRODUCTION

New medical studies are being published continuously and clinicians are faced with increasingly large amounts of new information, to the point where it has become nearly impossible for clinicians to read and evaluate all available data in a medical field. In addition, the study results are not consistently reproducible and the results of individual studies are often insufficient to provide confident answers [1]. Many studies have false-positive results (type I errors) or are unable to detect small effects (false-negative, type II errors) [2]. Consequently, clinical decision making is particularly difficult when the published data are conflicting or the sample size is too small to be reliable [3].

Evidence-based medicine is the best available evidence in the medical literature [4]. Moreover, the best evidence in evidence-based medicine is from meta-analyses, which provide a less biased, more precise estimate on a clinical issue [5]. "Meta" comes from the Greek for "after" or "beyond"; a meta-analysis is an "analysis of analyses" [6]. In other words, it is a statistical technique for combining the results from different studies on the same topic [7], and is becoming popular for resolving discrepancies in clinical research. As such, a meta-analysis is an objective, quantitative synthesis of research findings [7] that increases the statistical strength and precision for estimating effects by combining the results of previous studies and, thus, overcoming the problem of small sample sizes and inadequate statistical strength [8]. A meta-analysis can explore the sources of heterogeneity, and identify subgroups associated with the factor of interest, potentially providing new insights for future studies [9]. When conducted properly, a meta-analysis

Copyright © 2018 The Korean Association of Internal Medicine
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

pISSN 1226-3303
eISSN 2005-6648
http://www.kjim.org

SHOUT
SHARE
IT

of medical studies is considered decisive evidence, as it occupies a top level in the hierarchy of evidence [5]. Therefore, meta-analysis is a more efficient and effective standard method for summarizing the results of many studies than is subjective judgment; therefore, it has become an important research strategy, and is progressively expanding. For example, a PubMed search of the title word "meta-analysis" produced 120,537 possible articles.

It is necessary to understand the statistical principles of meta-analysis in evidence-based medical practice, and it is important that clinicians understand its methods, advantages, and limitations. Therefore, this review introduces the basic concepts, steps, advantages, and caveats of meta-analysis, to help clinicians understand meta-analysis in clinical practice and research.

HOW TO PERFORM A META-ANALYSIS

Table 1 summarizes the general process of meta-analysis in medical studies. General or specialized statistical programs, such as STATA, SAS, R, Review Manager, or Comprehensive Meta-Analysis, are used to perform the statistical analysis of a meta-analysis.

Identification of relevant studies

The identification of all relevant papers is a critical component in a meta-analysis because its outcome depends on the studies included [10]. The search strategy has to be comprehensive, and more than one database should be searched. The three electronic databases that are considered to be the most important sources of medical studies are commonly searched, i.e., PubMed, EMBASE, and The Cochrane Central Register of Controlled Trials (CENTRAL) [11].

Heterogeneity

A meta-analysis examines the existence of heterogeneity among primary studies and analyzes the variance in the results of different studies; meta-analysis heterogeneity is the degree of dissimilarity in the individual study results [7]. The heterogeneity test examines the null hypothesis, i.e., there are no differences in the findings of the primary studies. Statistical tests, such as Cochran's Q test and the I^2 value, have been developed to detect and quantify heterogeneity in meta-analysis. Specifically, Cochran's Q test is used to determine whether there are differences between primary studies or if the variation seen is due to chance [12]. Cochran's Q-value is calculated by summing the squared deviations of the estimate of each study from the overall estimate and subsequently comparing it with the chi-square distribution with $\kappa-1$ degrees of freedom (df), where κ is the number of studies [12]. However, the Q test may be unreliable when the meta-analysis involves a small number of studies. Therefore, a heterogeneity $p < 0.10$ (not 0.05) indicates the presence of heterogeneity, since Cochran's Q test has low statistical strength and is insensitive [13]. Another commonly used method for testing heterogeneity is the I^2 value, which quantifies the effect of heterogeneity, and does not depend on the number of studies or the type of outcome data. I^2 values range between 0% and 100%, and represent the proportion of inter-study variability that can be attributed to heterogeneity rather than chance [$I^2 = 100\% \times (Q - df)/Q$] [14]. I^2 values of 25%, 50%, and 75% are considered low, moderate, and high estimates, respectively [14,15].

Table 1. General method of medical studies meta-analysis

1. Identify relevant research: PubMed, EMBASE, CENTRAL[a]
2. Check between-study heterogeneity: Cochran Q test, I^2
3. Meta-analysis: fixed/random effect model, Forrest plot
4. Evaluate sources of the heterogeneity: subgroup analysis, sensitivity test, meta-regression
5. Check publication bias: funnel plot, Egger's regression test, trim and fill method
6. Present meta-analysis result based on PRISMA[b]

[a]CENTRAL: The Cochrane Central Register of Controlled Trials.
[b]PRISMA: preferred reporting items for systematic reviews and meta-analyses.

https://doi.org/10.3904/kjim.2016.195

Fixed versus random effects models

A meta-analysis combines the effect sizes of the included studies by weighting the data according to the different amounts of information in each study. The weights are calculated using the sample size or the variance of each study [8]. There are two statistical models for a meta-analysis: the fixed effect and random effect models. The fixed effect model assumes that all of the studies in the meta-analysis have one true effect size, and the observed variation among studies is caused by sampling errors or chance [16]. The random effect model assumes that different studies exhibit substantial diversity, and the true effect size may vary from study to study [17]. Consequently, the fixed effect model assesses only intra-study sampling errors (intra-study variation), while the random effect model assesses both intra-study sampling errors and inter-study variance (between-study variation) [18]. As such, the choice of meta-analysis model depends on the presence or absence of heterogeneity. In the absence of heterogeneity (heterogeneity $p \geq 0.10$), a fixed effect model is used. However, when the Q-value is significant ($p < 0.10$), indicating the existence of heterogeneity in the studies, a random effect model should be used for the meta-analysis [19]. When the study groups are homogeneous, both models offer similar results; nonetheless, in the case of heterogeneity, the random effect model typically provides wider confidence intervals (CIs) than the fixed effect one [20].

Evaluation of the causes of heterogeneity

It is important to assess the presence of heterogeneity among the studies included in a meta-analysis, and determine the possible causes of heterogeneity, since it can lead to bias, referred to as "mixing apples and oranges," in the meta-analysis results [21]. As a result, subgroup analysis is used to assess the impact of heterogeneity, using factors such as ethnicity, number of studies, or clinical features to assess the impact of a potential source of heterogeneity. Sensitivity testing may also be performed to assess the influence of each individual study on the pooled effect size by omitting each individual study [22]. Meta-regression is used to explore the reasons for the heterogeneity and adjust for confounding effects, and is feasible if sufficient data are reported in the individual studies [23].

Publication bias

Studies showing positive effects tend to be published more frequently than those that do not, and studies showing no significant results tend to remain unpublished [24]. As a meta-analysis includes only published studies, it might overestimate the actual effect degree [24]. This outcome is called "publication bias." In other words, a meta-analysis may be subject to publication bias. The funnel plot is a commonly used graphic test to assess publication bias in a meta-analysis [25]. This test is a scatterplot of the effect estimate from each study in the meta-analysis against the measure of its precision (1/standard error) or sample size [26]. The effect estimates of small studies will scatter at the bottom of the graph, while the spread of larger studies will be narrower. In the absence of publication bias, the funnel plot produces a symmetrical inverted funnel, asymmetry being suggestive of a publication bias [27]. The funnel plot is a simple method, but it is difficult to interpret when the number of studies is small and can be misleading. Due to the limitations of funnel plots, which require a range of studies of varying sizes, involving subjective judgments, publication bias can be evaluated using other methods, such as Egger's linear regression test [25], which measures funnel plot asymmetry using a natural logarithm scale of odds ratios. Egger's regression test examines whether the intercept deviates significantly from zero in a regression of the standardized effect estimates against their precision. When asymmetry is present, the "trim and fill" method adjusts the summary estimates for observed bias [28]. This method removes or adds small studies until funnel plot symmetry is achieved by recalculating the center of the funnel before removing studies and replacing them with their missing mirror-image counterparts. A revised summary estimate is subsequently calculated using all of the original studies and the hypothetical "filled" studies [28].

Meta-analysis publication

Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) is an evidence-based set of items developed for reporting systematic reviews and meta-analyses [29]. Following the PRISMA recommendations helps authors to improve meta-analysis reporting.

STRENGTHS AND WEAKNESSES OF META-ANALYSIS

A major advantage of a meta-analysis is that it produces a precise estimate of the effect size with considerably increased statistical power, which is especially important when the power of the primary study is limited because of the small sample size. A meta-analysis also analyzes the variation in the results of different studies and quantifies result inconsistency (heterogeneity) across studies. It is also an objective, quantitative method that provides a less biased estimate on a specific topic. The main criticism of meta-analysis is that it combines different types of studies ("mixing apples and oranges") [7]. Nonetheless, this problem can be overcome by assessing the heterogeneity in the studies and performing subgroup analysis [30]. However, if studies are too heterogeneous to be comparable, a meta-analysis should be avoided, as the meta-analysis result may be meaningless and any true effect may be obscured. The other limitation of meta-analysis is "garbage in, garbage out," which means that if a meta-analysis includes low-quality studies with bias, the results of the meta-analysis will be biased and incorrect [31]. As such, the results of the meta-analysis depend on the quality of the primary research. Therefore, any meta-analysis should include studies selected based on strict inclusion criteria.

EXAMPLE OF A META-ANALYTIC STUDY

Tofacitinib is a novel oral Janus kinase inhibitor [32]. Several clinical trials have attempted to evaluate the efficacy and safety of tofacitinib in active rheumatoid arthritis (RA). Therefore, a meta-analysis approach to randomized clinical trial (RCT) data was used in an attempt

Table 2. Meta-analysis of randomized controlled trials on the efficacy of tofacitinib in active RA

Tofacitinib dosedose	Outcome	No. of Studies	Test of association			Test of heterogeneity		
			WMD	95% CI	*p* value	Model	*p* value	I^2
Tofacitinib 5 mg, bid	ACR20	3	2.445[a]	1.229 to 4.861	0.011	R	0.014	76.7
	Tender joint count	3	−5.731	−8.054 to −3.048	1.3×10^{-7}	F	0.279	21.6
	Swollen joint count	3	−5.422	−9.593 to −1.252	0.011	R	0.008	79.4
	Pain (VAS)	3	−12.72	−18.06 to −7.376	3.0×10^{-7}	F	0.464	0
	Patient global assessment	3	−17.82	−28.20 to −7.444	$< 1.0 \times 10^{-8}$	R	0.017	75.6
	Physician global assessment	3	−17.88	−26.48 to −9.286	$< 1.0 \times 10^{-8}$	R	0.067	63.0
	HAQ	3	−0.341	−0.455 to −0.226	$< 1.0 \times 10^{-8}$	F	0.526	0
	CRP	3	−16.43	−28.09 to −4.778	0.006	R	0.000	87.8
Tofacitinib 10 mg, bid	ACR20	3	2.597[a]	1.514 to 4.455	0.001	R	0.054	65.8
	Tender joint count	3	−6.295	−8.517 to −4.073	2.0×10^{-9}	F	0.639	0
	Swollen joint count	3	−5.970	−9.630 to −2.311	0.001	R	0.010	78.5
	Pain (VAS)	3	−18.20	−29.50 to −8.230	0.002	R	0.020	74.5
	Patient global assessment	3	−17.70	−27.17 to −8.230	$< 1.0 \times 10^{-8}$	R	0.044	68.0
	Physician global assessment	3	−17.45	−28.82 to −6.082	0.003	R	0.010	78.2
	HAQ	3	−0.344	−0.461 to −0.227	$< 1.0 \times 10^{-8}$	F	0.143	48.5
	CRP	3	−17.07	−32.15 to −1.999	0.026	R	0.000	92.2

WMD, weighted mean difference; CI, confidence interval; bid, twice daily; ACR20, American College of Rheumatology 20% response rate; R, random effects model; F, fixed effects model; VAS, visual analog scale; HAQ, Health Assessment Questionnaire; CRP, C-reactive protein.
[a]Relative risk.

https://doi.org/10.3904/kjim.2016.195

to increase the precision and accuracy of estimates of the efficacy and safety of tofacitinib 5 and 10 mg twice daily in active RA. We performed a systematic review and meta-analysis of RCTs, examining the efficacy and safety of tofacitinib in active RA patients using PubMed, EMBASE, CENTRAL, and manual searches [33]. Five RCTs, including three phase 2 and two phase 3 trials and 1,590 patients, met the inclusion criteria [33]. The three phase 2 RCTs included 452 RA patients in the meta-analysis. The American College of Rheumatology 20% response rate (ACR20) response rate was significantly higher in the tofacitinib 5 mg group than in the controls (relative risk [RR], 2.445; 95% CI, 1.229 to 4.861; $p = 0.011$) (Table 2, Fig. 1). Similarly, the ACR20 response rate was significantly higher in the tofacitinib 10 mg group than in the controls (RR, 2.597; 95% CI, 1.514 to 4.455; $p = 0.001$) (Table 2, Fig. 1). Significant improvements were observed in the tofacitinib 5 and 10 mg groups compared with the controls for all efficacy outcomes, such as the number of tender and swollen joints, pain, patient and physician global assessments of disease activity, the Health Assessment Questionnaire, and C-reactive protein levels (Table 2). The safety outcomes did not differ between the tofacitinib 5 and 10 mg groups and placebo groups, except for infection in the tofacitinib 10 mg group (RR, 2.133; 95% CI, 1.268 to 3.590; $p = 0.004$) (Table 2, Fig. 2). The two phase 3 trials confirmed the findings of the meta-analysis of the phase 2 studies. The meta-analysis found that tofacitinib, at dosages 5 or 10 mg twice daily, was effective in active RA and had a manageable safety profile.

CONCLUSIONS

Meta-analyses in medical research cover a wide range of topics, from risk factors to prognosis [34,35]. Therefore, meta-analysis is applicable to a broad spectrum of topics, including biomarkers, genetic factors, diagnosis, and treatment [33,36-38]. It is also a powerful method for combining the results of different studies, and for summarizing current evidence on a specific issue objectively. Consequently, a meta-analysis provides a more precise estimate of effect sizes and investigates sources of variation and difference effects among subgroups. It can also resolve conflicts between studies, and yield con-

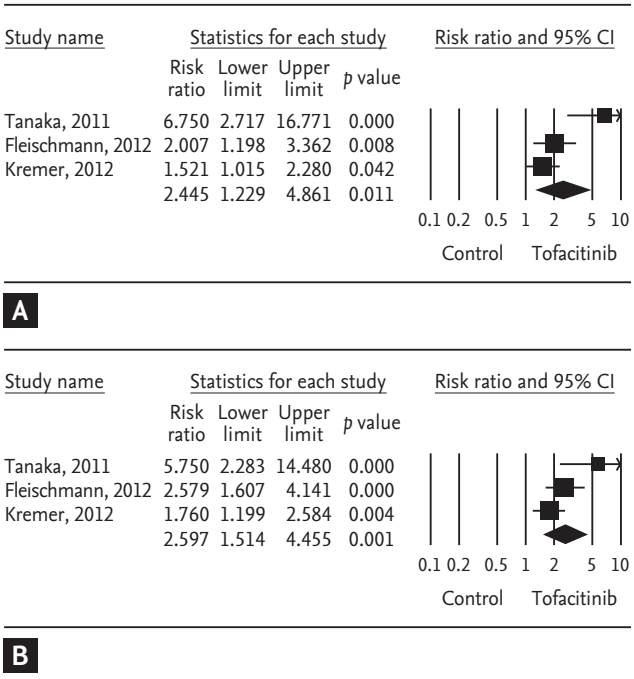

Figure 1. Meta-analysis of the efficacy of tofacitinib (A) 5 mg and (B) 10 mg twice a day on American College of Rheumatology 20% response rate in rheumatoid arthritis. CI, confidence interval.

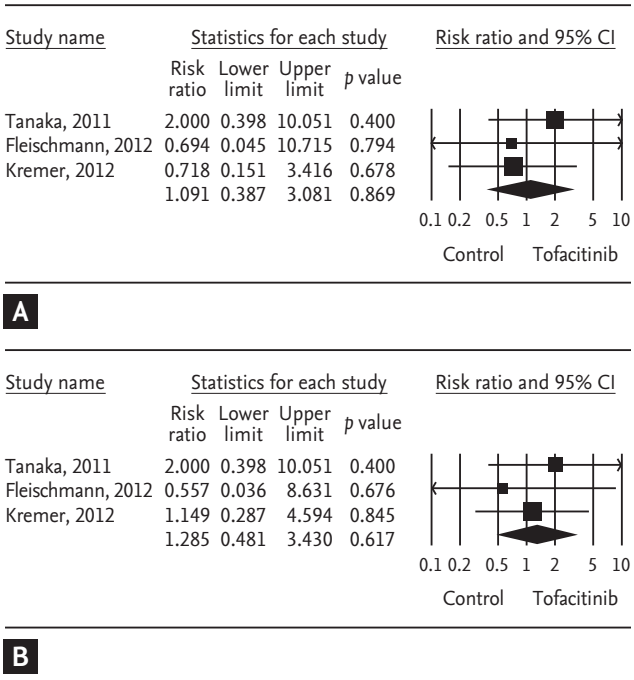

Figure 2. Meta-analysis of the efficacy of tofacitinib (A) 5 mg and (B) 10 mg twice a day on number of patients withdrawn due to adverse events in rheumatoid arthritis. CI, confidence interval.

clusive results when individual studies are inconclusive. Meta-analysis is an invaluable bridge between past and future studies. Therefore, understanding meta-analysis is valuable for clinicians in the process of making clinical decisions.

Conflict of interest

No potential conflict of interest relevant to this article was reported.

REFERENCES

1. Murad MH, Montori VM, Ioannidis JP, et al. How to read a systematic review and meta-analysis and apply the results to patient care: users' guides to the medical literature. JAMA 2014;312:171-179.
2. Button KS, Ioannidis JP, Mokrysz C, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013;14:365-376.
3. Haidich AB. Meta-analysis in medical research. Hippokratia 2010;14(Suppl 1):29-37.
4. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. Clin Orthop Relat Res 2007;455:3-5.
5. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. JAMA 1995;274:1800-1804.
6. Glass GV. Primary, secondary, and meta-analysis of research. Educ Res 1976;5:3-8.
7. Walker E, Hernandez AV, Kattan MW. Meta-analysis: Its strengths and limitations. Cleve Clin J Med 2008;75:431-439.
8. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. BMJ 1997;315:1533-1537.
9. Gotzsche PC. Why we need a broad perspective on meta-analysis. It may be crucially important for patients. BMJ 2000;321:585-586.
10. Leenaars M, Hooijmans CR, van Veggel N, et al. A step-by-step guide to systematically identify all relevant animal studies. Lab Anim 2012;46:24-31.
11. Qi X, Yang M, Ren W, et al. Find duplicates among the PubMed, EMBASE, and Cochrane Library Databases in systematic review. PLoS One 2013;8:e71838.
12. Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med 1991;10:1665-1677.
13. Munafo MR, Flint J. Meta-analysis of genetic association studies. Trends Genet 2004;20:439-444.
14. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-1558.
15. Thakkinstian A, McElduff P, D'Este C, Duffy D, Attia J. A method for meta-analysis of molecular association studies. Stat Med 2005;24:1291-1306.
16. Davey Smith G, Egger M. Meta-analyses of randomised controlled trials. Lancet 1997;350:1182.
17. DerSimonian R, Kacker R. Random-effects model for meta-analysis of clinical trials: an update. Contemp Clin Trials 2007;28:105-114.
18. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychol Methods 1998;3:486-504.
19. Ried K. Interpreting and understanding meta-analysis graphs: a practical guide. Aust Fam Physician 2006;35:635-638.
20. Zintzaras E, Lau J. Synthesis of genetic association studies for pertinent gene-disease associations requires appropriate methodological and statistical approaches. J Clin Epidemiol 2008;61:634-645.
21. Salanti G, Sanderson S, Higgins JP. Obstacles and opportunities in meta-analysis of genetic association studies. Genet Med 2005;7:13-20.
22. Copas J, Shi JQ. Meta-analysis, funnel plots and sensitivity analysis. Biostatistics 2000;1:247-262.
23. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1559-1573.
24. Dickersin K, Min YI. Publication bias: the problem that won't go away. Ann N Y Acad Sci 1993;703:135-146.
25. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-634.
26. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin Epidemiol 2001;54:1046-1055.
27. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. BMJ 2006;333:597-600.
28. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56:455-463.
29. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA

https://doi.org/10.3904/kjim.2016.195

Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009;151:264-269.
30. Bailar JC 3rd. The promise and problems of meta-analysis. N Engl J Med 1997;337:559-561.
31. Sharpe D. Of apples and oranges, file drawers and garbage: why validity issues in meta-analysis will not go away. Clin Psychol Rev 1997;17:881-901.
32. Changelian PS, Flanagan ME, Ball DJ, et al. Prevention of organ allograft rejection by a specific Janus kinase 3 inhibitor. Science 2003;302:875-878.
33. Song GG, Bae SC, Lee YH. Efficacy and safety of tofacitinib for active rheumatoid arthritis with an inadequate response to methotrexate or disease-modifying antirheumatic drugs: a meta-analysis of randomized controlled trials. Korean J Intern Med 2014;29:656-663.
34. Lee YH. Meta-analysis of genetic association studies. Ann Lab Med 2015;35:283-287.
35. Lee YH, Bae SC, Song GG. Gene-environmental interaction between smoking and shared epitope on the development of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: a meta-analysis. Int J Rheum Dis 2014;17:528-535.
36. Song G, Bae SC, Lee YH. Diagnostic accuracy of lung ultrasound for interstitial lung disease in patients with connective tissue diseases: a meta-analysis. Clin Exp Rheumatol 2016;34:11-16.
37. Song GG, Seo YH, Kim JH, Choi SJ, Ji JD, Lee YH. Association between TNF-α (-308 A/G, -238 A/G, -857 C/T) polymorphisms and responsiveness to TNF-α blockers in spondyloarthropathy, psoriasis and Crohn's disease: a meta-analysis. Pharmacogenomics 2015;16:1427-1437.
38. Lee YH, Nath SK. Systemic lupus erythematosus susceptibility loci defined by genome scan meta-analysis. Hum Genet 2005;118:434-443.